10-2004

Trans-Normal distribution: A flexible model for duration and event-time data

Zhenlin YANG
Singapore Management University, zlyang@smu.edu.sg

Citation
YANG, Zhenlin. Trans-Normal distribution: A flexible model for duration and event-time data. (2004). 1-9. Research Collection School Of Economics.
Available at: https://ink.library.smu.edu.sg/soe_research/2062

Trans-Normal Distribution: A Flexible Model for Duration and Event-Time Data

Article · January 2004

1 author:

Zhenlin Yang
Singapore Management University

Trans-Normal Distribution: A Flexible Model for Duration and Event-Time Data

Zhenlin Yang*

School of Economics and Social Sciences, Singapore Management University
469 Kukit Timah Road, Singapore 259756

October 1, 2004

Abstract

A new family of distributions, called the *trans-normal distribution* is introduced. Its basic properties are presented. A closely related family of distributions called *tran-t* distribution is also given along with its basic properties. Trans-normal regression is used as an illustration on the usefullness of the new distribution.

Keywords: Hazard function; Survivor function; Trans-normal; Trans-normal regression; Trans-*t*.

1 Introduction

In the analysis of economic durations, medical event-times, and engineering reliability data, one constantly faces the problem of choosing a suitable distribution from many such as lognormal, Weibull, gamma, inverse Gaussian, and Birbaum-Saunders. In a recent article, Yang (2001), followed the work of Hernandez and Johnson (1980), has demonstrated that the Box-Cox transformation-based distribution (Box and Cox, 1964) provides a general confidence interval for the median of a future obervation, which performs equivalently to the confidence intervals that are specifically designed for a given distribution, but is more robust against distribution misspecification.

*Corresponding author. Tel.: +65-6822-0852; fax: +65-6822-0833; *E-mail address:* zlyang@smu.edu.sg (Z. L. Yang)

In this article, we formalize this notion and introduce a general family of distributions called the trans-normal distribution. Some basic properties of this distribution are presented. It is shown that this new family of distributions embeds several well known distributions and sub-famillies of distributions. It is illustrated that this distribution can generate all kinds of shapes for the hazard function and hence should be very useful in analyzing the economic durations, medical event-times, and engineering reliability data. A closely related distribution called *trans-t* distribution is also introduced and properties presented. A brief discussion on the trans-normal regression along the same line as the Box-Cox regression is given to illustrate the usefulness of this new distribution.

2 The Trans-Normal Family

Definition 1. *A family of distributions is called the* trans-normal *family if the random variable Y is such that $h(Y) \sim N(\mu, \sigma^2)$ for some $h: B \to R$, a monotonic increasing and differentiable function with range R, the whole real line, and domain $B \subseteq R$.*

Clearly, the pdf of the trans-normal distribution has the form:

$$f(y; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}[h(y) - \mu]^2\right\} h'(y), \tag{1}$$

where $h'(y) = dh(y)/dy$. The domain B could be the positive half real line as in the case of duration or time-to-event data, a bounded interval as in the case of percentage or proportions, and the whole real line such as certain economic data. The trans-normal family is seen to be a very rich family. It contains popular distributions such as normal with $h(y) = y$, and lognormal with $h(y) = \log y$. It also covers several sub families such as the ξ-normal family (Saunders, 1974) where $h(y)$ satisfies $h(y) = -h(y^{-1}), y > 0$ and $\alpha^{-1}h(y/\beta) \sim N(0, 1)$, and the Box and Cox (1964) power family:

$$h(y) = \begin{cases} (y^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \log y, & \lambda = 0, \end{cases} \quad y > 0. \tag{2}$$

When $\lambda \neq 0$, the Box-Cox power transformation is bounded either below or above depends on whether λ is positive or negative. Thus, exact normality for $h(Y)$ is not possible. In practice, one often takes the normality as an approximation. To overcome this difficulty, Yang (2002) proposed a modified power transformation:

$$h(y) = \begin{cases} (y^{\lambda} - y^{-\lambda})/2\lambda, & \lambda > 0, \\ \log y, & \lambda = 0, \end{cases} \quad y > 0. \tag{3}$$

For nonnegative y, this function is one-to-one with its inverse $y = [\lambda h + (1 + \lambda^2 h^2)^{1/2}]^{1/\lambda}$. Note that the distributional family generated by modified power transformation generalizes the ξ-normal family of Saunders (1974). For $h : R \to R$, Yeo and Johnson (2000) provides a good example (see also the references therein). We now give some general theoretical properties of the trans-normal family.

Theorem 1. *Let $f(y) = f(y; \mu, \sigma)$ be the pdf of a trans-normal random variable Y defined in (3). Assume $h(y)$ is monotonically increasing with the first two derivatives $h'(y)$ and $h''(y)$ exist. Define $m(y) = h''(y)/h'^2(y) - \sigma^{-2}(h(y) - \mu)$. Then $f(y)$*

i) is a monotonic function of y if $m(y) = 0$ does not have a real root;

ii) is a unimodal pdf if $m(y) = 0$ has a unique real root in the interior of B;

iii) has two stationary points if $m(y) = 0$ has two real roots;

iv) is bimodal if $m(y) = 0$ has three real roots, etc.

Proof. Let $k(y) = \exp\{-[h(y) - \mu]^2/(2\sigma^2)\}$. Then, $f(y) \propto k(y)h'(y)$ and $f'(y) = k(y)h'^2(y)[h''(y)/h'^2(y) - \sigma^{-2}(h(y) - \mu)] = k(y)h'^2(y)m(y)$. Since the function $k(y)h'^2(y)$ is a positive function of y, how many times that $f'(y)$ changes its sign as y changes depends on how many real roots that $m(y) = 0$ has, which determines the behavior of $f(y)$. The results of the theorem thus follows.

Note that the case (i) in Theorem 1 rarely happens, case (ii) is the most typical case and it happens as long as $f(y)$ vanishes at both ends and $h''(y)/h'^2(y)$ is monotonic. The cases (iii) is also not common and (iv) can happen for certain special functions at certain parameter settings. Let Φ denote the CDF of the standard normal distribution. The survivor and hazard functions of the trans-normal distribution are, respectively,

$$S(t; \mu, \sigma) = 1 - \Phi\left[\frac{h(t) - \mu}{\sigma}\right], \quad \text{and} \tag{4}$$

$$r(t; \mu, \sigma) = \frac{f(t; \mu, \sigma)}{1 - \Phi[(h(t) - \mu)/\sigma]}. \tag{5}$$

To illustrate the versatility and usefulness of the trans-normal distribution, we pick a special modified power transformation, $h(y) = y^{.5} - y^{-.5}$, and plot the pdf, the survivor function (sf) and the hazard function (hf) for serval parameter configurations. From the plots summarized in Figure 1, we see that the pdf of this trans-normal distribution has all kinds of shapes: it can be nearly symmetric, bimodal, or very skewed depending whether σ is small, medium, or large relative to the mean of Y. When σ is small relative to the mean, the pdf has one bump at the center part; as σ increases, another bump shows up at the left of the center and as σ further increases, the first bump disappeared and the distribution becomes unimodal again. Figure 1 also exhibits serval shapes of hazard function, including the interesting 'bath-tub' shape, which has a popular engineering interpretation: first bump represents the 'burn-in' period, the center flat part represents the 'stable period' and the second bump represents the 'wear-out' period. Econometricians call this the U-shaped hazard (Kiefer, 1988) and some evidence for its existence is provided by Kennan (1985) from the analysis of the strike duration data. It is interesting to note that when σ is large, the hf has a sharp increase at the very beginning and then quickly becomes flat for a long period of 'time'. This exactly reflects the failure mechanisms of certain engineering systems and electronic components which are very fragile at the very beginning, but once stabilized, can last for a very long period of time.

3 Some Related Results

Like the transition from normal to t distribution and the transition from lognormal to log-t (Dahiya and Guttman, 1982), the trans-normal distribution can also be extended to give a trans-t distribution.

Definition 2. *A family of distributions is called the* trans-t *family with ν degrees of freedom and with parameters ξ and τ if the random variable T is such that $[h(T) - \xi]/\tau$ follows a t-distribution with ν degrees of freedom.*

This definition is a generalization of the log-t distribution given in Dahiya and Guttman (1982), where they showed that a log-t pdf is either a purely decreasing function or a function with two stationary points. It can be further shown that the log-t pdf is essentially

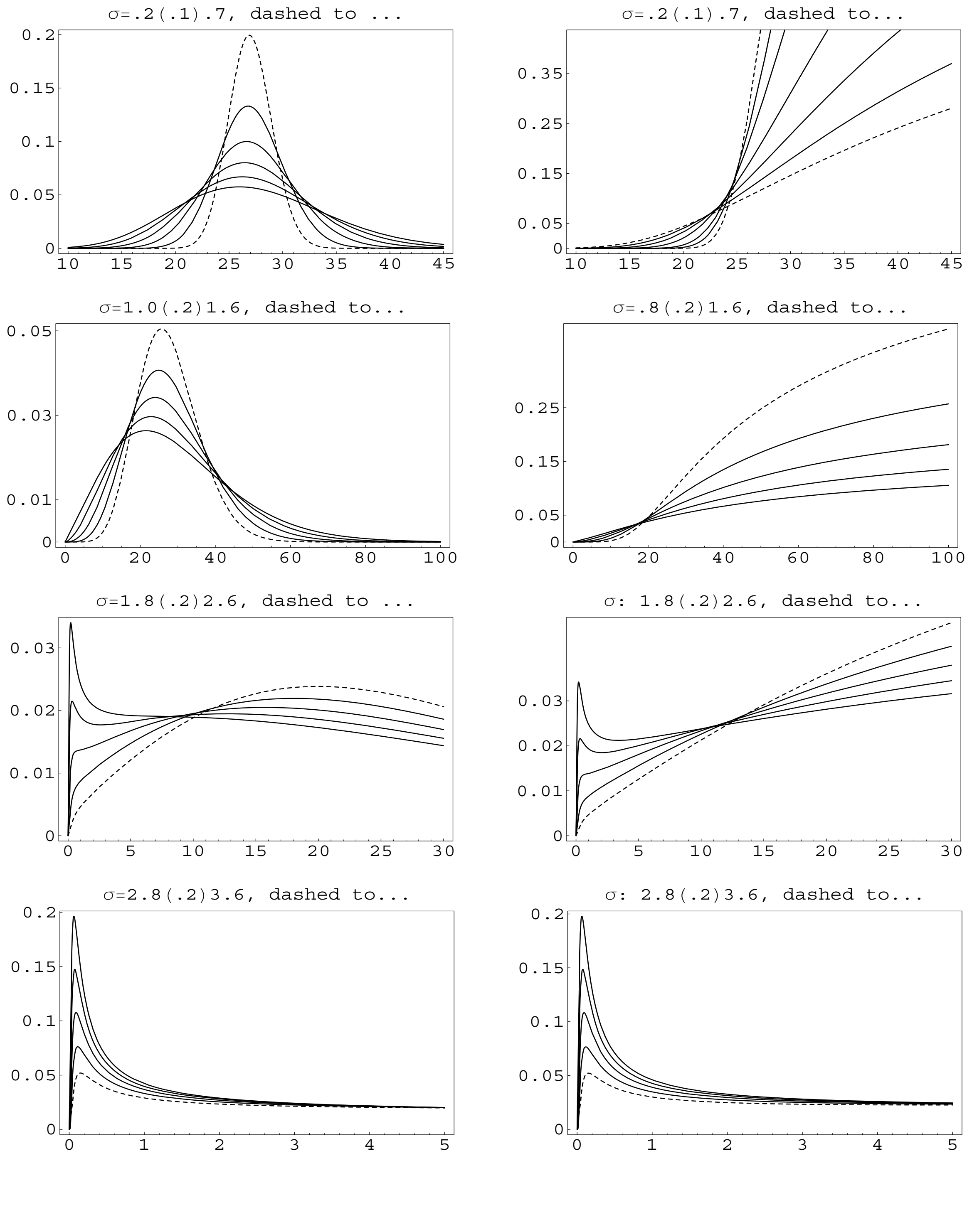

Figure 1: Plots of the trans-normal density function (left panel) and hazard function.

unimodal if the degrees of freedom ν is large relative to τ, which may often be the case in practical applications. We give some general properties of the trans-t in the following theorem. Let $g(t;\xi,\tau)$ be the pdf of a trans-t random variable. Then,

$$g(t;\xi,\tau) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\tau\Gamma(\nu/2)}\left[1+\frac{[h(t)-\xi]^2}{\nu\tau^2}\right]^{-\frac{\nu+1}{2}} h'(t). \tag{6}$$

Theorem 2. *Assume that the h function satisfies the conditions of Theorem 1. Define $m(t) = h''(t)/h'^2(t) - \{(\nu-1)[h(t)-\xi]\}/\{\nu\tau^2 + [h(t)-\xi]^2\}$. Then:*

(i) $g(t;\xi,\tau)$ is a decreasing function of t if $m(t)=0$ has no real root; is unimodal if $m(t)=0$ has a unique real root; has two stationary points if $m(t)=0$ has two real roots; bimodal if $m(t)=0$ has three real roots, etc.

(ii) The trans-t pdf converges to a trans-normal pdf as $\nu \to \infty$.

Proof. Define $k(t) = \{1+[h(t)-\xi]^2/(\nu\tau^2)\}^{-(\nu+1)/2}$, and let $g(t) = g(t;\xi,\tau)$ Then, we have $g(t) \propto k(t)h'(t)$, which goes to zero as $t \to \infty$. Also,

$$g'(t) = k(t)h'^2(t)\left[\frac{h''(t)}{h'^2(t)} - \frac{(\nu-1)[h(t)-\xi]}{\nu\tau^2+[h(t)-\xi]^2}\right] = k(t)h'^2(t)m(t).$$

As both $k(t)$ and $h'^2(t)$ are positive functions, how many times that $g'(t)$ changes its sign depends on the number of real roots of the equation $m(t)=0$. Hence the part (i) of the theorem follows. The proof of part (ii) is straightforward.

From the expression of $m(t)$, it is easy to see that when $h''(t)/h'^2(t)=0$, $m(t)$ has only one root, and when $h''(t)/h'^2(t) = const$, $m(t)$ has two roots. The former corresponds to the linear transformation that gives a unimodal pdf and the latter the log transformation that gives a pdf with two stationary points. Figure 2 gives an comparison between the trans-normal and trans-t distribution. It is seen that the trans-t pdf is generally more flatter than the corresponding trans-normal pdf.

The Trans-normal Regression. The trans-normal regression model has the form

$$h(\mathbf{Y}) = \mathbf{X}\beta + \varepsilon \tag{7}$$

where $h(\mathbf{Y})$ is an $n\times 1$ vector of the transformed responses, $\mathbf{X}$ is an $n \times p$ design matrix of full column rank, β is a $p\times 1$ vector of regression coefficients, and ε is the error vector

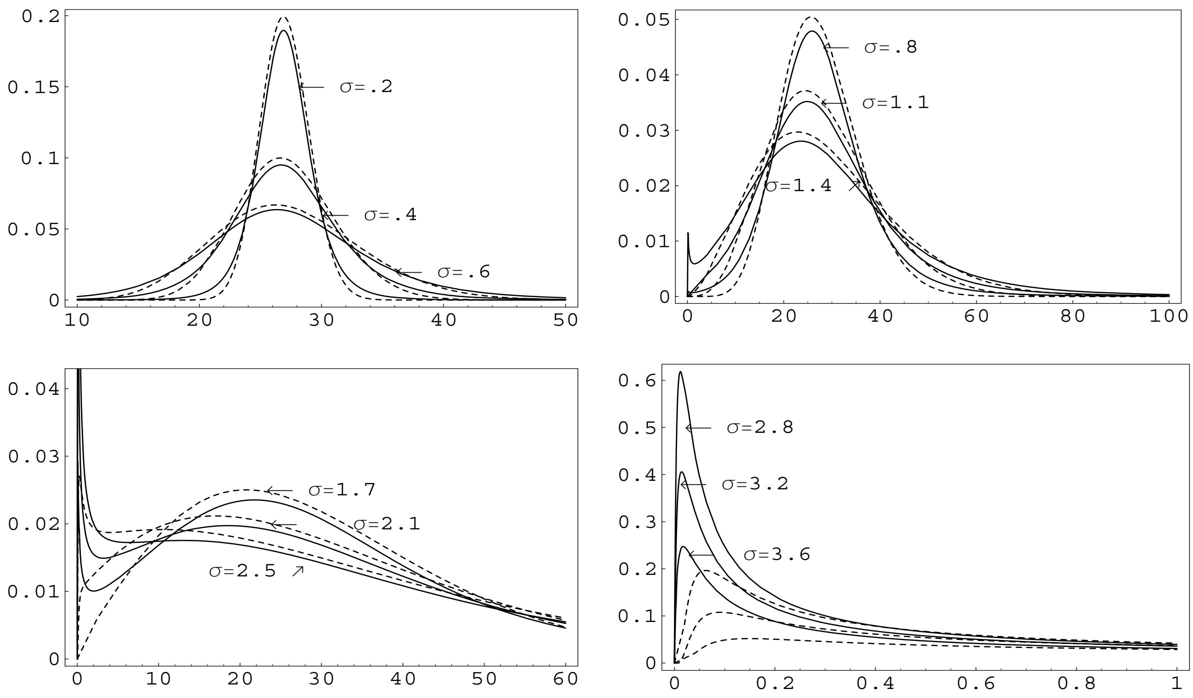

Figure 2: A Comparison of the Trans-normal density with the Trans-t Density

that is assumed to be $N(0_n, \sigma^2 I_n)$ with I_n being an $n \times n$ identity matrix. The model (7) is essentially the Box-Cox transformation model (Box and Cox, 1964). The difference is that h now is a proper transformation from B to R, and that it may be completely specified. In this sense, the exact normality for (7) is technically possible and hence one does not need to worry about the regularity conditions for the consistency and asymptotic normality of the parameter estimates.

The true survivor and hazard functions at a future value $\mathbf{x}_0'$ of the explanatory variables have the form $S(t;\beta,\sigma) = 1 - \Phi\left[(h(t) - \mathbf{x}_0'\beta)/\sigma\right]$ and $r(t;\beta,\sigma) = f(t;\beta,\sigma)/S(t;\beta,\sigma)$, respectively, where $f(t;\beta,\sigma)$ is the trans-normal pdf with parameters $\mathbf{x}_0'\beta$ and σ.

4 A Numerical Example

We now provide a numerical examples to illustrate the trans-normal distribution.

Example 1. A Set of Trans-normal Data. A sample of 68 observations are drawn from the model $h(Y, 0.25) \sim N(5, 1.5^2)$ with h being the power transformation: 39.00 25.38

17.08 79.04 13.65 4.80 19.22 40.55 68.56 18.16 21.53 23.08 8.80 21.05 28.34 29.09 34.49 23.67 12.15 12.02 33.83 15.43 56.59 14.93 21.50 12.69 2.01 42.66 46.74 42.42 8.23 16.51 29.75 21.54 31.69 26.25 26.28 9.74 57.97 25.68 98.70 13.71 11.50 33.82 41.69 33.97 38.65 20.34 12.70 16.52 14.03 24.64 10.15 40.15 6.05 42.09 34.45 42.93 57.06 14.01 19.00 26.11 19.69 33.25 29.12 19.26 41.83 24.06.

The MLEs of μ and σ with known $\lambda = 0.25$ are 4.8790 and 1.4079, respectively. The MLE of λ when it is assumed unknown is $\hat{\lambda} = 0.2863$, which gives the MLEs of μ and σ 5.2302 and 1.5772, respectively. Figure 2 gives plots of the estimated pdfs, sfs and hfs for the cases of λ known or unknown. The ML and FPD methods give similar estimates of the pdf and sf, but give substantially different estimates of hf. The FPD estimate of hf is much closer to the true function than the ML estimate, especially at medium and large durations. This is consistent with the simulation results given in the last section. From the plots, we do not see a clear effect of estimating transformation.

5 Conclusions

The applicability of the fiducial predictive density approach in the econometric duration analysis is studied. Two flexible families of duration distributions, the trans-normal and the trans-exponential, are proposed, for detailed examinations of this approach. The former has an easy extention to the trans-normal regression and the latter allows for the analysis of censored data. It is found that the FPD approach gives simple and reliable estimates of the density, survivor, and hazard functions, and that it provides a simple way to constructing the shortest prediction intervals. The results are further extended to the case that the transformation functional form is known, but the function is indexed by an unknown parameter. The latter extension of the results greatly expands the applicability of the FPD method in the econometric duration analysis.

Acknowlegements

Research support from the Wharton-SMU reseach center, Singapore Management University is gratefully acknowledged.

References

1. Box, G. E. P and Cox, D. R. (1964). An analysis of transformation (with discussion). *Journal of the Royal Statistical Soceity* B 26, 211-252.
2. Dahiya, R. C. and Guttman, I. (1982). Shortest confidence and prediction intervals for the log-normal. *Canadian Journal Statistics* 10, 277-291.
3. Hernandez F. and Johnson, R. A. (1980). The large sample behavior of transformations to normality. *Journal of the American Statistical Association* 75, 855-861.
4. Kennan, J. (1985). The duration of contract strikes in U.S. manufacturing. *Journal of Econometrics* 28, 5-28.
5. Kiefer, N. M. (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature* 26, 646-679.
6. Saunders, S. C. (1974). A family of random variables closed under reciprocation. *Journal of the American Statistical Association* 69, 533-539.
7. Yang, Z. L. (2001). Predicting a future median life through a power transformation. *Lifetime Data Analysis.* **7**, 305-317.
8. Yang, Z. (2002). A modified family of power transformations. *Working Paper*, School of Economics and Social Sciences, Singapore Management University.
9. Yeo, I. K. and Johnson, R. A. (2000). A new family of power transformation to improve normality or symmetry. *Biometrika*, **87**, 954-959.